

Swendsen-Wang is faster than single-bond dynamics

Mario Ullrich*

Mathematisches Institut, Universität Jena

email: mario.ullrich@uni-jena.de

January 30, 2012

Abstract

We prove that the spectral gap of the Swendsen-Wang dynamics for the random-cluster model is larger than the spectral gap of a single-bond dynamics, that updates only a single edge per step. For this we give a representation of the algorithms on the joint (Potts/random-cluster) model.

Furthermore we obtain upper and lower bounds on the mixing time of the single-bond dynamics on the discrete d -dimensional torus of side length L at the Potts transition temperature for q large enough that are exponential in L^{d-1} , complementing a result of Borgs, Chayes and Tetali [BCT10].

1 Introduction

This work was motivated by the recent article of Borgs, Chayes and Tetali “Tight Bounds for Mixing of the Swendsen-Wang Algorithm at the Potts Transition Point” [BCT10], where the authors prove upper and lower bounds for the mixing time of the Swendsen-Wang and heat-bath dynamics for the q -state Potts model on rectangular subsets (of side length L) of the lattice \mathbb{Z}^d with periodic boundary conditions at the Potts transition temperature if q and L are large enough. Both, upper and lower bounds, are exponential in L^{d-1} . (Their upper bounds are valid for all q, L .) Since one can sample from the Potts model if one can do so for the random-cluster model

*The author was supported by the DFG GK 1523.

(see e.g. [Gri06]), we wonder if the same upper and lower bounds are valid for the heat-bath dynamics for the random-cluster model, i.e. the *single-bond dynamics*.

In this article we give a positive answer to this question. For the lower bound we prove that if G be an arbitrary graph, $p \in (0, 1)$ and $q \in \mathbb{N}$, then

$$\lambda(P_{\text{SW}}) \geq \lambda(P_{\text{SB}}),$$

where $\lambda(\cdot)$ denotes the spectral gap and P_{SW} (resp. P_{SB}) denotes the transition matrix of the Swendsen-Wang (resp. single-bond) dynamics as defined in Section 4 (see Theorem 5). For this we represent both Markov chains by (transition) matrices on the joint (Potts/random-cluster) model. Estimates of the norm of the corresponding Markov operators lead to the result.

By this inequality (and [BCT10]) we obtain an exponential (in L^{d-1}) lower bound for the mixing time of the single-bond dynamics on boxes of side length L in \mathbb{Z}^d with periodic boundary condition at the transition point for q and L large enough, like for the Swendsen-Wang dynamics. The proof of the upper bound uses a result of Ge and Štefankovič [GŠ10] that provides us with an upper bound on the mixing time of the single-bond dynamics in terms of the *linear-width* of a graph (see Section 7).

Furthermore we obtain some rapid mixing results for the Swendsen-Wang dynamics. First we easily obtain rapid mixing of the Swendsen-Wang dynamics on trees (which was proven by several authors before, see e.g. [CF98], [BCT10]), because if $T = (V, E)$ is a tree, the random-cluster measure is a product measure and so the single-bond dynamics has spectral gap $\Omega(1/|E|)$ (see e.g. [LPW09, Lemma 12.11]). Hence, $\lambda(P_{\text{SW}})^{-1} = \mathcal{O}(|E|)$ for every tree T , $p \in (0, 1)$ and $q \in \mathbb{N}$, see Corollary 7. Additionally we get (again by [GŠ10]) that Swendsen-Wang is rapidly mixing for graphs with bounded linear-width (Corollary 8).

2 The models

Fix some $p \in (0, 1)$, a natural number $q \geq 1$ and a graph $G = (V, E)$ with finite vertex set V and edge set E .

The *random-cluster model* (also known as the FK-model), see Fortuin and Kasteleyn [FK72], is defined on the graph $G = (V, E)$ by its state space $\Omega_{\text{RC}} = \{A : A \subseteq E\}$ and the RC measure

$$\mu(A) := \mu_{p,q}^G(A) = \frac{1}{Z(G, p, q)} \left(\frac{p}{1-p} \right)^{|A|} q^{c(A)},$$

where $c(A)$ is the number of connected components in the graph (V, A) , counting isolated vertices as a component, and Z is the normalization constant that makes μ a probability measure. For a detailed introduction and related topics see [Gri06]. Although μ is well-defined for all $q > 0$, we are only interested in integer values. In this case there is a tight connection to a model on (proper and improper) colorings of the vertices of the graph G . The q -state Potts model on G at inverse temperature $\beta \geq 0$ is defined as the set of possible *configurations* $\Omega_P = [q]^V$, where $[q] := \{1, \dots, q\}$ is the set of *colors* (or spins), together with the probability measure

$$\pi(\sigma) := \pi_{\beta,q}^G(\sigma) = \frac{1}{Z(G, 1 - e^{-\beta}, q)} \exp \left\{ \beta \sum_{u,v: u \leftrightarrow v} \mathbb{1}(\sigma(u) = \sigma(v)) \right\}$$

for $\sigma \in \Omega_P$, where $u \leftrightarrow v$ if u and v are *neighbors* in G , i.e. $\{u, v\} \in E$, and $Z(\cdot, \cdot, \cdot)$ is the same normalization constant as for the RC model (see [Gri06, Th. 1.10]).

To describe the algorithms we also need the coupling of the Gibbs measure $\pi_{\beta,q}^G$ and the random-cluster measure $\mu_{p,q}^G$ of Edwards and Sokal [ES88]. Let us define

$$\Omega(A) := \{ \sigma \in \Omega_P : \sigma(u) = \sigma(v) \ \forall \{u, v\} \in A \}, \quad A \in \Omega_{RC},$$

and

$$E(\sigma) := \{ \{u, v\} \in E : \sigma(u) = \sigma(v) \}, \quad \sigma \in \Omega_P.$$

Obviously, we have for $\sigma \in \Omega_P$ and $A \subset E$ that $\sigma \in \Omega(A) \Leftrightarrow A \subset E(\sigma)$. Let $\sigma \in \Omega_P$, $A \in \Omega_{RC}$ and $p = 1 - e^{-\beta}$, then the joint measure of $(\sigma, A) \in \Omega_J := \Omega_P \times \Omega_{RC}$ is

$$\bar{\mu}(\sigma, A) := \bar{\mu}_{p,q}^G(\sigma, A) = \frac{1}{Z(G, p, q)} \left(\frac{p}{1-p} \right)^{|A|} \mathbb{1}(A \subset E(\sigma)).$$

The marginal distributions of $\bar{\mu}$ are exactly π and μ , respectively, and we will call $\bar{\mu}$ the *FKES* (Fortuin-Kasteleyn-Edwards-Sokal) *measure*.

3 Mixing time

In the following we want to estimate the efficiency of Markov chains. For an introduction to Markov chains and techniques to bound the convergence rate to the stationary distribution, see e.g. [LPW09]. Let P be the transition matrix of a Markov chain with state space Ω that is ergodic, i.e. irreducible and aperiodic, and

has unique stationary measure π . Then we define the *mixing time* of the Markov chain by

$$\tau(P) := \min \left\{ t : \max_{x \in \Omega} \sum_{y \in \Omega} |P^t(x, y) - \pi(y)| \leq \frac{1}{e} \right\}.$$

If we are considering simultaneously a *family* of state spaces $\{\Omega_n\}_{n \in \mathbb{N}}$ with a corresponding family of Markov chains $\{P_n\}_{n \in \mathbb{N}}$, we say that the chain is *rapidly mixing* for the given family if $\tau(P_n)^{-1} = \mathcal{O}(\log(|\Omega_n|)^C)$ for all $n \in \mathbb{N}$ and some $C < \infty$.

Another quantity that seems to be more convenient if we want to compare two Markov chains is the spectral gap. For this let the Markov chain P be additionally reversible with respect to π , i.e.

$$\pi(x) P(x, y) = \pi(y) P(y, x) \quad \text{for all } x, y \in \Omega.$$

(All the transition matrices from this article satisfy these condition.)

Then we know that the eigenvalues of P are real and we define the *spectral gap* by

$$\lambda(P) = 1 - \max \left\{ |\xi| : \xi \text{ is an eigenvalue of } P, \xi \neq 1 \right\}$$

The eigenvalues of the Markov chain can be expressed in terms of norms of the operator P that maps from $L_2(\pi) := (\mathbb{R}^\Omega, \pi)$ to $L_2(\pi)$, where inner product and norm are given by $\langle f, g \rangle_\pi = \sum_{x \in \Omega} f(x)g(x)\pi(x)$ and $\|f\|_\pi^2 := \sum_{x \in \Omega} f(x)^2\pi(x)$, respectively. The operator is defined by

$$Pf(x) := \sum_{y \in \Omega} P(x, y) f(y) \tag{1}$$

and represents the expected value of the function f after one step of the Markov chain starting in $x \in \Omega$. The *operator norm* of P is

$$\|P\|_\pi := \|P\|_{L_2(\pi) \rightarrow L_2(\pi)} = \max_{\|f\|_\pi \leq 1} \|Pf\|_\pi$$

and we use $\|\cdot\|_\pi$ interchangeably for functions and operators, because it will be clear from the context which norm is used. It is well known that $\lambda(P) = 1 - \|P - S_\pi\|_\pi$ for reversible P , where $S_\pi(x, y) = \pi(y)$. We know that reversible P are self-adjoint, i.e. $P = P^*$, where P^* is the (*adjoint*) operator that satisfies $\langle f, Pg \rangle_\pi = \langle P^*f, g \rangle_\pi$ for all $f, g \in L_2(\pi)$. The mixing time and spectral gap of a Markov chain (on finite state spaces) are closely related by the following inequality, see e.g. [LPW09, Theorem 12.3 & 12.4].

Lemma 1. *Let P be the transition matrix of a reversible, ergodic Markov chain with state space Ω and stationary distribution π . Then*

$$\lambda(P)^{-1} - 1 \leq \tau(P) \leq \log\left(\frac{2e}{\pi_{\min}}\right) \lambda(P)^{-1},$$

where $\pi_{\min} := \min_{x \in \Omega} \pi(x)$.

4 The algorithms

The *Swendsen-Wang dynamics* (on the RC model) is based on the given connection of the random cluster and Potts models and performs the following two steps:

- 1) Given a random cluster state $A_t \in \Omega_{\text{RC}}$ on G , assign a random color independently to each connected component of (V, A) . Vertices of the same component get the same color. This gives $\sigma \in \Omega_{\text{P}}$.
- 2) Take $E(\sigma)$ and delete each edge independently with probability $1 - p$. This gives $A_{t+1} \in \Omega_{\text{RC}}$.

This can be seen as first choosing σ with respect to the conditional probability of $\bar{\mu}$ given A_t and then choosing A_{t+1} with respect to $\bar{\mu}$ given σ . The transition matrix of the Swendsen-Wang dynamics is given by

$$P_{\text{SW}}(A, B) = q^{-c(A)} \left(\frac{p}{1-p}\right)^{|B|} \sum_{\sigma \in \Omega_{\text{P}}} (1-p)^{|E(\sigma)|} \mathbb{1}(\sigma \in \Omega(A \cup B)). \quad (2)$$

Note that one can consider the Swendsen-Wang dynamics also on the Potts model (as it is usually done), which, starting at some $\sigma \in \Omega_{\text{P}}$, performs the two steps above in reverse order. It is easy to prove that Swendsen-Wang on Potts and random-cluster model have the same spectral gap, see e.g. [U11, Sec. 2.4].

The second algorithm we want to analyze is the (lazy) *single-bond dynamics*. Let $A \in \Omega_{\text{RC}}$ be given and denote by $\overset{A}{\leftrightarrow}$ (resp. \nleftrightarrow^A) connected (resp. not connected) in the subgraph (V, A) . Additionally we use throughout this article $A \cup e$ instead of $A \cup \{e\}$ (respectively for \cap, \setminus). This chain performs the following steps:

- 1) With probability $\frac{1}{2}$ set $B = A$. Otherwise, choose an edge $e = \{e_1, e_2\} \in E$ uniformly at random.

- 2) (i) If $e_1 \xleftrightarrow{A} e_2$:
- $B = A \cup e$ with probability p .
 - $B = A \setminus e$ with probability $1 - p$.
- (ii) If $e_1 \xleftrightarrow{A} e_2$:
- $B = A \cup e$ with probability $\frac{p}{q}$.
 - $B = A \setminus e$ with probability $1 - \frac{p}{q}$.

3) The new state is B .

The transition matrix of this Markov chain can be written as

$$P_{\text{SB}}(A, B) = \frac{I(A, B)}{2} + \frac{1}{2|E|} \sum_{e \in E} P_e(A, B), \quad (3)$$

where $I(A, B) = \mathbb{1}(A = B)$ and P_e is given by

$$P_e(A, B) = \mathbb{1}(A \ominus B \subset e) \cdot \begin{cases} p^{|B \cap e|} (1 - p)^{1 - |B \cap e|}, & e_1 \xleftrightarrow{A} e_2 \\ \left(\frac{p}{q}\right)^{|B \cap e|} \left(1 - \frac{p}{q}\right)^{1 - |B \cap e|}, & e_1 \xleftrightarrow{A} e_2. \end{cases}$$

Here, \ominus denotes the symmetric difference.

Remark 2. If one is interested in the usual heat-bath dynamics on the random-cluster model, i.e.

$$\tilde{P}(A, B) := \frac{1}{2|E|} \sum_{e \in E} \frac{\mu(B)}{\mu(A \cup e) + \mu(A \setminus e)} \mathbb{1}(A \ominus B \subset e) \quad \text{for } A \neq B \quad (4)$$

and $\tilde{P}(A, A)$ such that \tilde{P} is stochastic, then all results of this article hold up to a constant. This is because $P_{\text{SB}}(A, B) \leq \tilde{P}(A, B) \leq (1 - p(1 - q^{-1}))^{-1} P_{\text{SB}}(A, B)$ for all $A \neq B$ and so it is easy to prove by standard techniques (see e.g. [DSC93]) that

$$\lambda(P_{\text{SB}}) \leq \lambda(\tilde{P}) \leq (1 - p(1 - q^{-1}))^{-1} \lambda(P_{\text{SB}}).$$

5 Representation on the joint model

In this section we want to represent the Swendsen-Wang and the single-bond dynamics on the FKES model, which consists of the product state space $\Omega_J := \Omega_P \times \Omega_{RC}$ and the FKES measure $\bar{\mu}$. For this we need the following “building blocks”. First we introduce the stochastic matrix that defines the mapping (by matrix multiplication) from the RC to the FKES model

$$M(B, (\sigma, A)) := q^{-c(B)} \mathbb{1}(A = B) \mathbb{1}(\sigma \in \Omega(B)). \quad (5)$$

Note that M defines an operator (like in (1)) that maps from $L_2(\bar{\mu})$ to $L_2(\mu)$ and its adjoint operator M^* can be given by the (stochastic) matrix

$$M^*((\sigma, A), B) = \mathbb{1}(A = B).$$

The following matrix represents the updates of the RC “coordinate” in the FKES model. For $(\sigma, A), (\tau, B) \in \Omega_J$ and $e = \{e_1, e_2\} \in E$ let

$$T_e((\sigma, A), (\tau, B)) := \mathbb{1}(\sigma = \tau) \begin{cases} p, & B = A \cup e \text{ and } \sigma(e_1) = \sigma(e_2) \\ 1 - p, & B = A \setminus e \text{ and } \sigma(e_1) = \sigma(e_2) \\ 1, & B = A \setminus e \text{ and } \sigma(e_1) \neq \sigma(e_2). \end{cases} \quad (6)$$

Before we state the Swendsen-Wang and the single-bond dynamics in terms of the matrices from (5) and (6), we state some properties that will be useful.

Lemma 3. *Let M , M^* and T_e be the matrices from above. Then*

- (i) M^*M and T_e are self-adjoint in $L_2(\bar{\mu})$.
- (ii) $M M^*(A, B) = \mathbb{1}(A = B)$ and thus $M^*M M^*M = M^*M$.
- (iii) $T_e T_e = T_e$ and $T_e T_{e'} = T_{e'} T_e$ for all $e, e' \in E$.
- (iv) $\|T_e\|_{\bar{\mu}} = 1$ and $\|M^*M\|_{\bar{\mu}} = 1$.

Proof. Part (i) and (ii) follow from the definition. Part (iii) comes from the fact that the transition probabilities depend only on the “coordinate” that will not be changed and (iv) follows from (i), (ii) and (iii), because $\|T_e\|_{\bar{\mu}} \stackrel{(iii)}{=} \|T_e^2\|_{\bar{\mu}} \stackrel{(i)}{=} \|T_e\|_{\bar{\mu}}^2$. \square

Now we can state the desired Markov chains with the matrices from above.

Lemma 4. *Let M , M^* and T_e be the matrices from above. Then*

$$(i) \quad P_{\text{SW}} = M \left(\prod_{e \in E} T_e \right) M^*.$$

$$(ii) \quad P_{\text{SB}} = \frac{I}{2} + \frac{1}{2|E|} \sum_{e \in E} M T_e M^*.$$

From Lemma 3(iii) we have that the order of multiplication in (i) is unimportant.

Proof. For (i) note that

$$\left(\prod_{e \in E} T_e \right) ((\sigma, A), (\tau, B)) = \mathbb{1}(\sigma = \tau) \mathbb{1}(B \subset E(\sigma)) p^{|B|} (1-p)^{|E(\sigma)|-|B|}.$$

Hence,

$$\begin{aligned} M \left(\prod_{e \in E} T_e \right) M^*(A, B) &= \sum_{\sigma \in \Omega_{\text{P}}} M(A, (\sigma, B)) \left(\prod_{e \in E} T_e \right) ((\sigma, A), (\tau, B)) \\ &= \sum_{\sigma \in \Omega_{\text{P}}} q^{-c(A)} \mathbb{1}(\sigma \in \Omega(A) \cap \Omega(B)) p^{|B|} (1-p)^{|E(\sigma)|-|B|} \\ &= P_{\text{SW}}(A, B). \end{aligned}$$

For (ii) it is enough to prove $P_e = M T_e M^*$, where P_e is from (3). First we define $\mathbb{1}_e(\sigma) := \mathbb{1}(\sigma(e_1) = \sigma(e_2))$ and $\mathbb{1}_e(A) := \mathbb{1}(e_1 \xleftrightarrow{A} e_2)$ for $\sigma \in \Omega_{\text{P}}$, $A \in \Omega_{\text{RC}}$ and $e = \{e_1, e_2\} \in E$. Now write

$$T_e((\sigma, A), (\sigma, B)) = \mathbb{1}(B = A \setminus e) + p \mathbb{1}_e(\sigma) \left[\mathbb{1}(B = A \cup e) - \mathbb{1}(B = A \setminus e) \right]$$

and note that $|\Omega(A)| = q^{c(A)}$ and

$$q^{-c(A)} \sum_{\sigma \in \Omega(A)} \mathbb{1}_e(\sigma) = \frac{1}{q} + \mathbb{1}_e(A) \left(1 - \frac{1}{q} \right).$$

Hence,

$$\begin{aligned}
MT_e M^*(A, B) &= \sum_{\sigma} q^{-c(A)} \mathbf{1}(\sigma \in \Omega(A)) T_e((\sigma, A), (\sigma, B)) \\
&= \mathbf{1}(B = A \setminus e) + p \left[\mathbf{1}(B = A \cup e) - \mathbf{1}(B = A \setminus e) \right] \\
&\quad \cdot q^{-c(A)} \sum_{\sigma \in \Omega(A)} \mathbf{1}_e(\sigma) \\
&= \begin{cases} p, & B = A \cup e \text{ and } e_1 \overset{A}{\leftrightarrow} e_2 \\ 1 - p, & B = A \setminus e \text{ and } e_1 \overset{A}{\leftrightarrow} e_2 \\ \frac{p}{q}, & B = A \cup e \text{ and } e_1 \overset{A}{\nleftrightarrow} e_2 \\ 1 - \frac{p}{q}, & B = A \setminus e \text{ and } e_1 \overset{A}{\nleftrightarrow} e_2. \end{cases} \\
&= P_e(A, B) \quad \text{for all } A, B \in \Omega_{\text{RC}} \text{ with } A \ominus B \subset e.
\end{aligned}$$

□

6 Main result

In this section we prove the following theorem.

Theorem 5. *Let P_{SW} and P_{SB} be the transition matrices of the Swendsen-Wang and single-bond dynamics from (2) and (3), respectively. Then*

$$\lambda(P_{\text{SW}}) \geq \lambda(P_{\text{SB}}).$$

This holds for arbitrary graphs G , $p \in (0, 1)$ and $q \in \mathbb{N}$.

Before we prove the theorem we state some corollaries. The first one gives an analogous inequality for the mixing times of the two algorithms.

Corollary 6. *Let P_{SW} and P_{SB} be the transition matrices of the Swendsen-Wang and single-bond dynamics for the random-cluster model on $G = (V, E)$ with parameters $p \in (0, 1)$ and $q \in \mathbb{N}$. Then*

$$\tau(P_{\text{SW}}) \leq \left(3 + |E| \log \frac{1}{p(1-p)} + |V| \log q \right) \tau(P_{\text{SB}}).$$

Proof. By Lemma 1 and Theorem 5 we obtain

$$\begin{aligned}\tau(P_{\text{SW}}) &\leq \log\left(\frac{2e}{\mu_{\min}}\right) \lambda(P_{\text{SW}})^{-1} \leq \log\left(\frac{2e}{\mu_{\min}}\right) \lambda(P_{\text{SB}})^{-1} \\ &\leq \log\left(\frac{2e}{\mu_{\min}}\right) \tau(P_{\text{SB}}) + 1 \\ &\leq (3 + \log(\mu_{\min}^{-1})) \tau(P_{\text{SB}}).\end{aligned}$$

(Note for the last inequality that $\tau(P_{\text{SB}}) = 0$ iff $\tau(P) = 0$ for every P .)
Since μ_{\min}^{-1} can easily be bounded by $(p(1-p))^{-|E|} q^{|V|}$ the result follows. \square

The next two corollaries show some rapid mixing results for the Swendsen-Wang dynamics. These are stated in terms of the spectral gap, but by Lemma 1 one can also use mixing times, loosing the same factor as in Corollary 6. The first one is rapid mixing if the underlying graph is a tree, which is already known (see e.g. [CF98], [BCT10]). The second is rapid mixing for graphs with bounded linear-width, which follows from a result of Ge and Štefankovič [GŠ10]. For this we define the *linear-width* of a graph $G = (V, E)$ as the smallest number ℓ such that there exists an ordering $e_1, \dots, e_{|E|}$ of the edges with the property that for every $i \in [|E|]$ there are at most ℓ vertices that have an adjacent edge in $\{e_1, \dots, e_i\}$ and in $\{e_{i+1}, \dots, e_{|E|}\}$. See [GŠ10] for bounds on the linear-width of paths, cycles, trees and in terms of a related quantity, the tree-width.

Corollary 7. *Let P_{SW} be the transition matrix of the Swendsen-Wang dynamics for the random-cluster model on a tree $T = (V, E)$. Then*

$$\lambda(P_{\text{SW}})^{-1} \leq 2(1 - p(1 - q^{-1}))^{-1} |E|.$$

Proof. Since $\mu_{p,q}^T$ is a product measure and we can write

$$P_{\text{SB}}(A, B) = \frac{1}{|E|} \sum_{e \in E} \left(\frac{I + P_e}{2} \right) (A \cap e, B \cap e) \prod_{f \neq e} \mathbb{1}(A \cap f = B \cap f),$$

where $(I + P_e)/2$ can be seen (here) as a 2×2 -matrix, i.e. the transition matrix of the single-bond dynamics on a single edge. This matrix has the eigenvalues 1 and $(1 + p(1 - q^{-1}))/2$. We obtain by [LPW09, Lemma 12.11] that $\lambda(P_{\text{SB}})^{-1} = 2(1 - p(1 - q^{-1}))^{-1} |E|$. This concludes the proof. \square

Note that this bound improves the one given in [BCT10, Corollary 3.2], because it does not depend on maximum degree and depth of the tree. The next results follows immediately from [GŠ10].

Corollary 8. *Let P_{SW} and P_{SB} be the transition matrices of the Swendsen-Wang and single-bond dynamics for the random-cluster model on a graph $G = (V, E)$ with linear-width bounded by ℓ . Then*

$$\lambda(P_{\text{SW}})^{-1} \leq \lambda(P_{\text{SB}})^{-1} \leq 4|E|^2 q^{\ell+1}. \quad (7)$$

Proof. The first inequality is Theorem 5 and the second follows from [GŠ10]. In this article the authors consider the Metropolis version of the single-bond dynamics. This Markov chain has transition probabilities

$$P_{\text{M}}(A, A \ominus e) = \frac{1}{2|E|} \min \left\{ 1, q^{c(A \ominus e) - c(A)} \left(\frac{p}{1-p} \right)^{|A \ominus e| - |A|} \right\}, \quad A \subset E,$$

with $P_{\text{M}}(A, A)$ such that P_{M} is a stochastic matrix. For this Markov chain they proof a lower bound on the congestion, which is defined as follows. Let $\Gamma = \{\gamma_{AB} : A, B \subset E\}$ where γ_{AB} are paths from A to B in the (directed) graph $\mathcal{H} = (\Omega_{\text{RC}}, \mathcal{E})$ with $\mathcal{E} = \{(A, B) : P_{\text{M}}(A, B) > 0\}$. Then we define the *congestion* of P_{M} (w.r.t. Γ) by

$$\varrho(P_{\text{M}}, \Gamma) := \max_{(B_1, B_2) \in \mathcal{E}} \frac{1}{\mu(B_1) P_{\text{M}}(B_1, B_2)} \sum_{A, C : (B_1, B_2) \in \gamma_{AC}} |\gamma_{AC}| \mu(A) \mu(C),$$

where $|\gamma_{AC}|$ denotes the length of the path. The bound of [GŠ10, Lemma 16] is $\varrho(P_{\text{M}}, \Gamma) \leq 2|E|^2 q^\ell$ for a suitable choice of Γ and so we obtain by [DS91, Prop. 1] (note that P_{M} is lazy) that

$$\lambda(P_{\text{M}})^{-1} \leq 2|E|^2 q^\ell.$$

But since it is easy to show that $P_{\text{M}}(A, B) \leq 2q P_{\text{SB}}(A, B)$ for all $A, B \subset E$, we can conclude by standard techniques (see e.g. [DSC93, eq. (2.3)]) that

$$\lambda(P_{\text{SB}})^{-1} \leq 2q \lambda(P_{\text{M}})^{-1} \leq 4|E|^2 q^{\ell+1}.$$

□

6.1 Proof of Theorem 5

First we need the following technical lemma.

Lemma 9. *Let H and G be two Hilbert spaces with corresponding inner products $\langle f, f' \rangle_H$ and $\langle g, g' \rangle_G$ for $f, f' \in H$ and $g, g' \in G$. Furthermore, let $A : G \rightarrow H$ be a bounded linear operator with adjoint operator A^* , i.e. $\langle A^*f, g \rangle_G = \langle f, Ag \rangle_H$ for all $f \in H$, $g \in G$, and let $B : G \rightarrow G$ be a positive (i.e. $\langle Bg, g \rangle_G \geq 0$), bounded, self-adjoint linear operator, then*

$$\|ABBA^*\|_{H \rightarrow H} \leq \|B\|_{G \rightarrow G} \|ABA^*\|_{H \rightarrow H}.$$

Proof. By the assumptions, B has a unique positive square root $B^{\frac{1}{2}}$, i.e. $B = B^{\frac{1}{2}}B^{\frac{1}{2}}$, which is again self-adjoint, see e.g. [Kre78, Th. 9.4-2]. We obtain

$$\begin{aligned} \|ABBA^*\|_{H \rightarrow H} &= \|AB\|_{G \rightarrow H}^2 \leq \|AB^{\frac{1}{2}}\|_{G \rightarrow H}^2 \|B^{\frac{1}{2}}\|_{G \rightarrow G}^2 \\ &= \|AB^{\frac{1}{2}}(AB^{\frac{1}{2}})^*\|_{H \rightarrow H} \|B^{\frac{1}{2}}(B^{\frac{1}{2}})^*\|_{G \rightarrow G} \\ &= \|ABA^*\|_{H \rightarrow H} \|B\|_{G \rightarrow G}. \end{aligned}$$

□

Now we are able to state the proof of Theorem 5.

Proof of Theorem 5. Let $S_\mu(A, B) := \mu(B)$ and $S_{(\mu, \bar{\mu})}(B, (\sigma, A)) = \bar{\mu}(\sigma, A)$ for all $A, B \in \Omega_{\text{RC}}$, $(\sigma, A) \in \Omega_{\text{J}}$. It is easy to verify that $S_{(\mu, \bar{\mu})}S_{(\mu, \bar{\mu})}^* = S_\mu$. Additionally we get from Lemma 3(iii) that $\prod_{e \in E} T_e = T_{e'} \prod_{e \in E} T_e$ for all $e' \in E$ and so

$$\prod_{e \in E} T_e = \prod_{e \in E} T_e \prod_{f \in E} T_f \quad (8)$$

and

$$\prod_{e \in E} T_e = \left(\frac{J}{2} + \frac{1}{2|E|} \sum_{f \in E} T_f \right) \prod_{e \in E} T_e, \quad (9)$$

where $J((\sigma, A), (\tau, B)) := \mathbb{1}((\sigma, A) = (\tau, B))$. Let $N := M - S_{(\mu, \bar{\mu})}$, then

$$\begin{aligned} \|P_{\text{SW}} - S_\mu\|_\mu &\stackrel{L.4}{=} \left\| M \left(\prod_{e \in E} T_e \right) M^* - S_\mu \right\|_\mu = \left\| N \left(\prod_{e \in E} T_e \right) N^* \right\|_\mu \\ &\stackrel{(8)}{=} \left\| N \left(\prod_{e \in E} T_e \right) \left(\prod_{f \in E} T_f \right) N^* \right\|_\mu = \left\| N \prod_{e \in E} T_e \right\|_{L_2(\bar{\mu}) \rightarrow L_2(\mu)}^2, \end{aligned}$$

because $N \prod_{e \in E} T_e$ induces an operator that maps from $L_2(\bar{\mu})$ to $L_2(\mu)$ and the operator $\prod_{g \in E} T_g N^*$ is its adjoint. Using submultiplicativity we obtain

$$\begin{aligned}
\|P_{\text{SW}} - S_\mu\|_\mu &\stackrel{(9)}{=} \left\| N \left(\frac{J}{2} + \frac{1}{2|E|} \sum_{f \in E} T_f \right) \prod_{e \in E} T_e \right\|_{L_2(\bar{\mu}) \rightarrow L_2(\mu)}^2 \\
&\leq \left\| N \left(\frac{J}{2} + \frac{1}{2|E|} \sum_{f \in E} T_f \right) \right\|_{L_2(\bar{\mu}) \rightarrow L_2(\mu)}^2 \left\| \prod_{e \in E} T_e \right\|_{\bar{\mu}}^2 \\
&\stackrel{L.3}{=} \left\| N \left(\frac{J}{2} + \frac{1}{2|E|} \sum_{f \in E} T_f \right) \right\|_{L_2(\bar{\mu}) \rightarrow L_2(\mu)}^2 \\
&= \left\| N \left(\frac{J}{2} + \frac{1}{2|E|} \sum_{f \in E} T_f \right)^2 N^* \right\|_\mu \\
&\stackrel{L.9}{\leq} \left\| M \left(\frac{J}{2} + \frac{1}{2|E|} \sum_{f \in E} T_f \right) M^* - S_\mu \right\|_\mu \stackrel{L.3}{=} \|P_{\text{SB}} - S_\mu\|_\mu,
\end{aligned}$$

where the last inequality comes from Lemma 9 with $H = L_2(\mu)$, $G = L_2(\bar{\mu})$, $B = \frac{J}{2} + \frac{1}{2|E|} \sum_{f \in E} T_f$ and $A = N$. Note that B is positive semidefinite (as stochastic matrix with $B(x, x) \geq \frac{1}{2}$ for all $x \in \Omega_J$) and $\|B\|_{\bar{\mu}} = 1$. This proves the claim, because $\lambda(P_{\text{SW}}) = 1 - \|P_{\text{SW}} - S_\mu\|_\mu$. \square

Remark 10. Note that the proof of Theorem 5 would be correct also in the case of the non-lazy single-bond dynamics, i.e. $P_{\text{SB}} = \frac{1}{|E|} \sum_{e \in E} P_e$, because similar to (9)

we have $\prod_{e \in E} T_e = (\frac{1}{|E|} \sum_{f \in E} T_f) \prod_{e \in E} T_e$ and $\frac{1}{|E|} \sum_{f \in E} T_f$ is positive semidefinite since so are all T_e . But for convenience of the proof of Corollary 8 we choose to consider the lazy version.

7 Mixing time bounds on the torus

In this section we prove that the mixing time of the single-bond dynamics on the discrete d -dimensional torus of side length L at the transition temperature is exponential in L^{d-1} , complementing a result of Borgs, Chayes and Tetali [BCT10]. For

the upper bound we use the bound of Corollary 8 together with a bound of the linear-width of the d -dimensional torus. The lower bound follows from the lower bound of the mixing time of the Swendsen-Wang dynamics from [BCT10, Theorem 1.2]. Since this is closely related (and also uses) their results, we refer to [BCT10] and the references cited therein for details. Let

$$T_{L,d} = (\mathbb{Z}/L\mathbb{Z})^d$$

be the d -dimensional torus of side length L . We will prove the following theorems.

Theorem 11. *Let P_{SB} be the single-bond dynamics for the random-cluster model on $T_{L,d}$ with parameters $p \in (0, 1)$ and $q \in \mathbb{N}$. Then for all $L, d \geq 2$ we have*

$$\tau(P_{\text{SB}}) \leq \exp\left\{k_1(p) + k_2(q) L^{d-1}\right\},$$

where

$$k_1(p) := \log\left(1 + \log\frac{1}{p(1-p)}\right)$$

and

$$k_2(q) := 4 + 3 \log q + \log(1 + \log q).$$

Theorem 12. *Let $d \geq 2$. Then there exists a constant $k_3 = k_3(d) > 0$ such that, for q and L sufficiently large, the single-bond dynamics for the random-cluster model on $T_{L,d}$ satisfies*

$$\tau(P_{\text{SB}}) \geq \exp\left\{k_3 \beta_0 L^{d-1}\right\} \quad \text{for } p = 1 - e^{\beta_0},$$

where β_0 is the Potts transition temperature, i.e. $\beta_0 = \frac{1}{d} \log q + \mathcal{O}(q^{-1/d})$.

First we prove Theorem 12.

Proof. For the proof we have to consider the Swendsen-Wang dynamics for the Potts model with Gibbs measure $\pi = \pi_{\beta,q}^{T_{L,d}}$ for $\beta = \beta_0$, that performs the two steps of the Swendsen-Wang dynamics (as given in Section 4) in reverse order. We denote its transition matrix by \tilde{P}_{SW} . These two algorithms have the same spectral gap if $p = 1 - e^{-\beta}$, see [U11, Sec. 2.4]. We obtain from Lemma 1 that

$$\tau(P_{\text{SB}}) + 1 \geq \lambda(P_{\text{SB}})^{-1} \stackrel{\text{Th.5}}{\geq} \lambda(P_{\text{SW}})^{-1} = \lambda(\tilde{P}_{\text{SW}})^{-1} \geq \log\left(\frac{2e}{\pi_{\min}}\right)^{-1} \tau(\tilde{P}_{\text{SW}}),$$

where $\pi_{\min} = \min_{\sigma \in \Omega_P} \pi(\sigma)$. Obviously, $\pi_{\min} \geq e^{-\beta|E|} q^{-|V|}$ for graphs $G = (V, E)$. We know from Theorem 1.2 of [BCT10] that there exists a constant $k'_3 = k'_3(d) > 0$ such that, for q and L large enough,

$$\tau(\tilde{P}_{\text{SW}}) \geq \exp\left\{k'_3 \beta_0 L^{d-1}\right\}, \quad \text{for } \beta = \beta_0.$$

Thus, since $|V| = L^d$ and $|E| = dL^d$ for $G = T_{L,d}$, we obtain

$$\tau(P_{\text{SB}}) + 1 \geq (2 + L^d \log q + \beta d L^d)^{-1} \exp\left\{k'_3 \beta_0 L^{d-1}\right\},$$

which implies (again for L large enough), that there exists a constant $k_3 = k_3(d) > 0$, such that

$$\tau(P_{\text{SB}}) \geq \exp\left\{k_3 \beta_0 L^{d-1}\right\} \quad \text{for } p = 1 - e^{\beta_0}.$$

□

For the proof of Theorem 11 we need the following lemma.

Lemma 13. *The linear-width of $T_{L,d}$ do not exceed $2L^{d-1} + 1$.*

Proof. For the proof we need 3 other “widths” of graphs, i.e. *path-width*, *proper path-width* and *bandwidth*, but we omit their definition, because we do not need them here. First note the following 3 facts:

1. linear-width is not larger than path-width+1 [FT05, Lemma 2]
2. path-width is not larger than proper path-width [KS96]
3. proper path-width equals bandwidth [KS96, Theorem 3.2]

Therefore it is enough to prove that $\text{bw}(T_{L,d})$, i.e. the bandwidth of $T_{L,d}$, is at most $2L^{d-1}$. For this note that $T_{L,d}$ is the cartesian product of d cycles $T_{L,1}$ of length L and that $\text{bw}(T_{L,1}) = 2$ (see [CCDG82, Theorem 4.1.1]). So we obtain by Corollary 4.3.2 of [CCDG82] that $\text{bw}(T_{L,d}) \leq 2L^{d-1}$. □

Now we are able to prove Theorem 11.

Proof of Theorem 11. Let l be the linear-width of $T_{L,d}$. We know from Lemma 13 that $l \leq 2L^{d-1} + 1$ and so $l + 1 \leq 3L^{d-1}$ since $L, d \geq 2$. It follows from Corollary 8 that

$$\lambda(P_{\text{SB}})^{-1} \leq 4d^2 L^{2d} q^{3L^{d-1}}.$$

Set $\eta = \frac{1}{p(1-p)}$ and using Lemma 1 we obtain

$$\begin{aligned}
\tau(P_{\text{SB}}) &\leq \log\left(\frac{2e}{\mu_{\min}}\right) \lambda(P_{\text{SB}})^{-1} \\
&\leq (2 + L^d \log q + dL^d \log \eta) 4d^2 L^{2d} q^{3L^{d-1}} \\
&\leq 4d^3 L^{3d} q^{3L^{d-1}} (1 + \log q + \log \eta) \\
&= \exp\left\{\log(4d^3 L^{3d}) + 3\log(q)L^{d-1} + \log(1 + \log q + \log \eta)\right\} \\
&\leq \exp\left\{4L^{d-1} + 3\log(q)L^{d-1} + \log(1 + \log q) + \log(1 + \log \eta)\right\} \\
&\leq \exp\left\{k_1(p) + k_2(q)L^{d-1}\right\}
\end{aligned}$$

with k_1 and k_2 from Theorem 11. This proves the claim. \square

References

- [BCT10] C. Borgs, J. Chayes, and P. Tetali, *Tight bounds for mixing of the Swendsen-Wang algorithm at the Potts transition point*, Probab. Theory Related Fields (2010), 1–49.
- [CCDG82] P. Z. Chinn, J. Chvátalov, A. K. Dewdney, and N. E. Gibbs, *The bandwidth problem for graphs and matrices - a survey*, J. Graph Theory **6** (1982), 223–254.
- [CF98] C. Cooper and A. M. Frieze, *Mixing properties of the Swendsen-Wang process on classes of graphs*, in proceedings of DIMACS workshop on statistical physics methods in discrete probability, combinatorics and theoretical computer science **41** (1998), 1499–1527.
- [DS91] P. Diaconis and D. Stroock, *Geometric bounds for eigenvalues of Markov chains*, Ann. Appl. Probab. **1** (1991), 36–61.
- [DSC93] P. Diaconis and L. Saloff-Coste, *Comparison theorems for reversible markov chains*, Ann. Appl. Probab. **3** (1993), 696–730.
- [ES88] S.F. Edwards and A.D. Sokal, *Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm*, Phys. Rev. D **38** (1988), 2009–2012.

- [FK72] C. M. Fortuin and P. W. Kasteleyn, *On the random-cluster model. I. Introduction and relation to other models*, Physica **57** (1972), 536–564.
- [FT05] F. V. Fomin and D. M. Thilikos, *A 3-approximation for the pathwidth of Halin graphs*, J. Discrete Algorithms **4** (2005), 499–510.
- [Gri06] G. Grimmett, *The random-cluster model*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 333, Springer-Verlag, Berlin, 2006.
- [GŠ10] Q. Ge and D. Štefankovič, *A graph polynomial for independent sets of bipartite graphs*, IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2010) (Dagstuhl, Germany) (Kamal Lodaya and Meena Mahajan, eds.), Leibniz International Proceedings in Informatics (LIPIcs), vol. 8, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2010, pp. 240–250.
- [Kre78] E. Kreyszig, *Introductory functional analysis with applications*, John Wiley & Sons, New York-London-Sydney, 1978.
- [KS96] H. Kaplan and R. Shamir, *Pathwidth, bandwidth and completion problems to proper interval graphs with small cliques*, SIAM J. Comput. **25** (1996), 540–561.
- [LPW09] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2009, With a chapter by J. G. Propp and D. B. Wilson.
- [U11] M. Ullrich, *Comparison of Swendsen-Wang and Heat-Bath Dynamics*, ArXiv e-prints (2011).